BIODIVERSITAS
Volume 25, Number 12, December 2024
Pages: 4860-4869

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d251222

Whole genome sequencing of a siderophore-producing bacterium isolated from Simadu pineapple roots

CUCU WIJAYANTI[1], PRAYOGA SURYADARMA[1,2,3,♥], NISA RACHMANIA MUBARIK[1,4]

[1]Program of Biotechnology, Graduate School, Institut Pertanian Bogor. Jl. Raya Dramaga, Kampus IPB Dramaga, Bogor 16680, West Java, Indonesia
[2]Department of Agroindustrial Technology, Institut Pertanian Bogor. Jl. Kamper, Kampus IPB Dramaga, Bogor 16680, West Java, Indonesia
[3]Biotech Center of Institut Pertanian Bogor. Jl. Raya Dramaga, Kampus IPB Darmaga, Bogor 16680, West Java, Indonesia. Tel.: +62-251-8621257, ♥email: prayoga@apps.ipb.ac.id
[4]Department of Biology, Institut Pertanian Bogor. Jl. Agatis, Kampus IPB Dramaga, Bogor 16680, West Java, Indonesia

Manuscript received: 2 August 2024. Revision accepted: 12 December 2024.

Abstract. *Wijayanti C, Suryadarma P, Mubarik NR. 2024. Whole genome sequencing of a siderophore-producing bacterium isolated from Simadu pineapple roots. Biodiversitas 25: 4860-4869.* Siderophores are organic compounds synthesized by bacteria that act as iron-chelating ligands, facilitating the solubilization and transport of iron into the cell. Siderophore-producing bacteria offer significant benefits across various fields of research and application. A recent study investigated a siderophore-producing bacterium, the M7 isolate, from Simadu pineapple roots in the Subang District of West Java, Indonesia. This study aimed to analyze the siderophore biosynthesis gene cluster (BGC) from the complete genome sequence of the M7 isolate through genome mining. Comparative genomic analysis with the strain *Providencia manganoxydans* LLDRA6 revealed that the M7 isolate has a similar guanine-cytosine (GC) content with a longer sequence length totaling 4,447,159 bp. Genome comparison using average nucleotide identity (ANI) calculations indicated that the M7 isolate had a genome similarity of 98.82% with strain *P. manganoxydans* LLDRA6. Additionally, the genome of the M7 isolate contains systems for iron acquisition and metabolism, which involve mechanisms for regulating iron. A specific region in the M7 isolate encodes the IucA/IucC family of siderophore biosynthesis proteins responsible for aerobactin biosynthesis through non-ribosomal peptide synthetase-independent siderophore (NIS) pathways. Regarding siderophore BGCs, this isolate showed the highest similarity to *Providencia stuartii*, with differences in genetic redundancy content. Additionally, pathogen prediction results for the M7 isolate indicate that it is not pathogenic to humans, unlike *P. stuartii*, which is known to be pathogenic. The new siderophore-producing bacterium, the M7 isolate, harbors a duplication of core biosynthesis genes, representing genetic redundancy that supports siderophore biosynthesis. This genetic feature may enable concurrent protein expression, potentially enhancing siderophore accumulation.

Keywords: Biosynthesis gene cluster, genome sequencing, *Providencia manganoxydans*, *P. stuartii*, siderophore-producing bacteria

Abbreviations: ANI: Average Nucleotide Identity; BGCs: Biosynthesis Gene Clusters; bp: base pair; GC: Guanine-Cytosine; NIS: Non-ribosomal peptide synthetase independent siderophore

INTRODUCTION

Iron is a vital micronutrient necessary for the survival of microorganisms. It plays a pivotal role in various biochemical processes, including electron transfer, metabolism, amino acid synthesis, and photosynthesis (Chandrangsu et al. 2017; Bradley et al. 2020; Seyoum et al. 2021). Microorganisms have evolved various strategies to obtain iron from their environment. One effective iron acquisition mechanism is the siderophore-dependent iron acquisition system, which facilitates iron uptake from the environment. Siderophores are small organic compounds synthesized under iron-limited conditions to bind and transport iron into microbial cells (Saha et al. 2016). Siderophores are classified into three main types based on their functional ligands: catecholate, hydroxamate, and carboxylate (Wilson et al. 2016; Swayambhu et al. 2021). Two pathways are involved in siderophore biosynthesis: non-ribosomal peptide synthetase (NRPS) dependent and NRPS-independent pathways (NIS). The NRPS pathway is significant in synthesizing catecholate siderophores, while the NIS pathway predominantly synthesizes hydroxamate and carboxylate siderophores (Miethke and Marahiel 2007).

Siderophores have diverse applications in agriculture, bioremediation, medicine, pharmacy, and biosensing technologies (Saha et al. 2016; Swayambhu et al. 2021). Their metal-binding properties make them useful in treating human diseases. For example, hydroxamate siderophores like desferrioxamine are used in Trojan horse antibiotics, Parkinson's disease treatment, and breast cancer therapies (Telfer et al. 2017; Bajbouj et al. 2018). Aerobactin, another hydroxamate siderophore, is valuable as a diagnostic tool in medical applications (Ho et al. 2017). These applications underscore the broad potential of siderophores in advancing therapeutic and diagnostic technologies.

Whole genome sequencing (WGS) of bacteria has become a commonly employed technique for analyzing bacterial genomes, allowing the identification of biosynthesis gene clusters (BGCs), such as those responsible for siderophore production, through genome mining (Cavas and Kirkiz 2022; Gu et al. 2024; Tamang et al. 2024). WGS aids in understanding the genetic basis of siderophore

production and discovering new ones with potential applications in medicine or agriculture. Bioinformatics tools like antiSMASH are crucial for identifying BGCs related to secondary metabolites, offering detailed and rapid analyses (Medema et al. 2011; Weber et al. 2015; Blin et al. 2021; Blin et al. 2023). Researchers can swiftly pinpoint and study BGCs that may produce compounds with antibiotic, antifungal, or anticancer properties. Moreover, integrating antiSMASH with other bioinformatics tools and databases enhances its predictive accuracy and broadens its application scope. WGS and bioinformatics have become essential for harnessing microbial diversity in biotechnological innovations. Identifying and characterizing siderophore BGCs shows great potential for enhancing iron acquisition in crops, developing novel antimicrobial agents, and advancing bioremediation efforts.

Bacterial species from genera like *Azotobacter*, *Bacillus*, *Pseudomonas*, *Streptomyces*, *Achromobacter*, *Klebsiella*, *Serratia, Achromobacter, Aeromonas, Alcaligenes, Burkholderia*, *Ochrobactrum, Stenotrophomonas* and *Providencia* are known to produce siderophores (Drechsel et al. 1993; Baars et al. 2016; Saha et al. 2016; Sasirekha and Srividya 2016; Yu et al. 2017; Majewska et al. 2024). Our previous research isolated a bacterium, the M7 isolate, from Simadu pineapple roots, known for producing sweeter fruit than other pineapples from Subang, Indonesia. Molecular identification based on 16S rRNA showed 87% similarity to *Providencia vermicola*, suggesting that M7 represents a different species within the *Providencia* genus, as similarity below 97% indicates species distinction (Bosshard et al. 2003; Fuadi et al. 2022). However, several *Providencia* species, including *Providencia vermicola*, *Providencia stuartii*, *Providencia rettgeri*, and *Providencia alcalifaciens* are pathogenic to humans and are linked to significant morbidity and mortality (Zavascki et al. 2015; Abdallah and Balshi 2018; Ameer and Al-Gburi 2022; Klein et al. 2024).

Therefore, it is crucial to assess the pathogenicity of the M7 isolate by examining the presence of genes associated with human pathogenicity in its genome to ensure its safety and potential benefits for various applications. While the 16S rRNA sequence only provides a general overview of phylogeny, it does not offer comprehensive details on overall bacterial genetics. This isolate has been identified as a siderophore-producing bacterium, with siderophore accumulation nearly 4-fold higher than *P. vermicola* after 12 hours (Fuadi et al. 2022), suggesting differences in the genes involved in siderophore biosynthesis. In addition, the M7 isolate produces a hydroxamate-type siderophore, indicating potential medical applications (Fuadi et al. 2022). This study aims to elucidate the genetic composition of the M7 isolate through complete genome analysis, identify BGCs involved in siderophore production, and determine its phylogenetic relationship with closely related species. A thorough investigation of its complete genome and associated biosynthesis gene clusters is essential to understand its molecular characteristics and potential uses fully.

MATERIALS AND METHODS

Bacterial strain

The M7 isolate was isolated from the roots of the Simadu pineapple (*Ananas comosus*) in the Subang district of West Java, Indonesia. Five g of roots were washed and soaked in sterile water, then immersed in 70% alcohol and rinsed three times with sterile water. The roots were ground in a mortar, and 1 mL of the homogenate was mixed with 9 mL of sterile distilled water. A 0.1 mL aliquot of the bacterial suspension, diluted to 10^{-2}-10^{-6}, was plated onto nutrient agar (NA) medium using the spread plate method. The plates were incubated at 28°C for 24 hours, and the colonies were periodically purified using the streak plate method (Fuadi et al. 2022). The M7 isolate was stored as frozen stocks in Luria Bertani (LB) medium containing 25% v/v glycerol. The strain is stored at 4°C in a NA medium for routine growth and maintenance.

Extraction of bacterial genomic DNA

Genomic DNA was extracted from an overnight culture grown in LB medium at 28°C with shaking at 120 rpm using a Quick-DNA HMW MagBead Kit (Zymo Research D6060, US) according to the manufacturer's protocol. The purity and concentration of genomic DNA were determined using a Nanodrop spectrophotometer and Qubit fluorometer (Paredes et al. 2021).

Whole genome sequencing

According to Nanopore template preparation protocols, WGS was performed using the GridION sequencer (Nanopore, Oxford, UK). GridION sequencing was operated using MinKNOW software. In the form of FastQ files generated from base calling with Guppy in high accuracy mode, the raw data were then subjected to quality control (QC) using Nanoplot (De Coster et al. 2018). De novo assembly was performed using Flye, initially corrected with Canu (Kolmogorov et al. 2020). The assembled sequences were polished four times with Racon (Vaser et al. 2017) and three times with Medaka (https://github.com/nanoporetech/medaka). The quality of the assembled sequence was assessed using Quast (Gurevich et al. 2013) and the Benchmarking Universal Single-Copy Orthologs (BUSCO) method (Simao et al. 2015). Annotation was carried out using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP), Rapid Annotation using Subsystem Technology (RAST), and visualization using Circos (Cui et al. 2020). Virulence gene analysis for the M7 isolate was conducted using VirulenceFinder 2.0 (Joensen et al. 2014). Additionally, pathogenicity predictions were made using PathogenFinder 1.1 (Cosentino et al. 2013).

Comparison genome analyses

The Average Nucleotide Identity (ANI) was calculated using the orthology (OrthoANI) algorithm in the OAT program, comparing the M7 isolate with nine other *Providencia* genomes downloaded from the GenBank database (https://www.ncbi.nlm.nih.gov) under accession GCA_016618195.1, GCA_014652175.1, GCA_000314855.2,

GCA_000314895.2, GCA_019048105.1, GCA_000156395.1, GCA_008693805.1, GCA_010748935.1, GCA_033802285.1.

Analysis of siderophore gene clusters

The antiSMASH 7.0 tool, in its bacteria version (https://antismash.secondarymetabolites.org/), was employed to comprehensively annotate BGCs involved in siderophore biosynthesis across the entire genome (Blin et al. 2023). Identified gene clusters were cross-referenced within the MIBiG repository. Default parameters were applied for antiSMASH analysis, with a relaxed detection strictness. The siderophore biosynthesis gene clusters were analyzed in other genome bacteria to compare the siderophore biosynthesis gene clusters with the M7 isolate.

RESULTS AND DISCUSSION

Genomic characteristics of the M7 isolate

The M7 isolate was subjected to whole genome sequencing using the Nanopore platform, providing a comprehensive overview of its genomic structure. The initial raw data, presented in FastQ file format, exhibited the sequencing characteristics of the M7 isolate, resulting in a substantial total of 500,000,920 bp in reads (Table 1). The raw sequencing data, including the N50 value of 3,951 bp, represents the median length of contigs. This data indicates that at least 50% of the assembled genome is covered by contigs of this length or longer. This metric is often used as an indicator of genome assembly quality. During the de novo assembly process, the reads were assembled into a single contig with a length of approximately 4,447,159 bp, reflecting the final assembled genome size. Additionally, the high genome coverage 185x indicates that nearly the entire genome has been sequenced and comprehensively obtained. The GC content of the M7 isolate's genome was 40.18%. This percentage reflects the guanine (G) and cytosine (C) bases in the genome and is often used to assess genomic characteristics such as genome stability or environmental adaptation. The GC content of 40.18% indicates that approximately 40% of the total bases in this genome are G or C. This value aligns with the typical GC content range for bacteria and can provide insights into this isolate's biological or ecological traits.

The BUSCO method assessed the completeness of conserved genes in the assembled genome. The results of the BUSCO evaluation of the M7 isolate indicated a gene completeness level of 98%, with 431 complete genes identified out of the total analyzed. The characteristics of the M7 isolate genome suggest that the genome assembly was performed effectively, resulting in a complete genome with consistent GC content and a very high level of gene completeness. The high-quality genome assembly of the M7 isolate will be valuable for subsequent phylogenetic analysis, allowing for detailed comparisons with other related strains and species.

Genome annotation of the M7 isolate

Based on the genome annotation to detect the functional elements within the previously arranged genome sequences, the annotated genomic data of the M7 isolate, covering various functional categories, was obtained using RAST. The results from RAST indicate that the genome of the M7 isolate possesses 3,827 subsystem features (Figure 1), including iron acquisition and metabolism, which involve mechanisms for regulating iron. The M7 isolate contains 12 genes contributing to the iron acquisition and metabolism system, including three genes related to siderophores and nine other genes involved in mechanisms that may be more complex and not yet fully understood. This result underscores the complexity of iron regulation in the M7 isolate's genome, highlighting the diverse genetic elements involved in this process. The importance of genome mining in this context cannot be overstated. Genome mining enables identifying and analyzing biosynthesis gene clusters associated with secondary metabolite production, such as siderophores, within the M7 isolate's genome. The genome mining techniques can pinpoint specific genes responsible for the biosynthesis of siderophores and other metabolites in the M7 isolate.

The genome annotation results, visualized using a Circos plot, provide detailed insights into the genomic features of the M7 isolate (Figure 2). The total number of genes is 4,060, with 3,948 coding sequences (CDSs), including those related to siderophore production. The 3,825 CDSs are associated with protein-coding genes, highlighting the genome's functional potential. The genome also comprises 112 RNA genes, including 8 5S rRNAs, 7 16S rRNAs, and 7 23S rRNAs, all complete. Additionally, there are 78 tRNAs and 12 non-coding RNAs (ncRNAs), essential for gene regulation and various cellular processes. The Circos plot effectively visualizes these features, depicting genetic variations such as shifts, duplications, and deletion and offering a comprehensive view of the genome's structural and functional complexity.

Table 1. Characteristics of the M7 isolate genome

Genome feature	Value
Number of bases	500,000,920 bp
N50	3,951 bp
Number of contigs	1
Largest contigs	4,447,159 bp
Mean read quality	14.0
Median read quality	14.1
Genome coverage	185x
GC content	40.18%
Complete and single-copy BUSCOs	428 (97.3%)
Complete and duplicated BUSCOs	3 (0.7%)
Fragmented BUSCOs	8
Missing BUSCOs	1

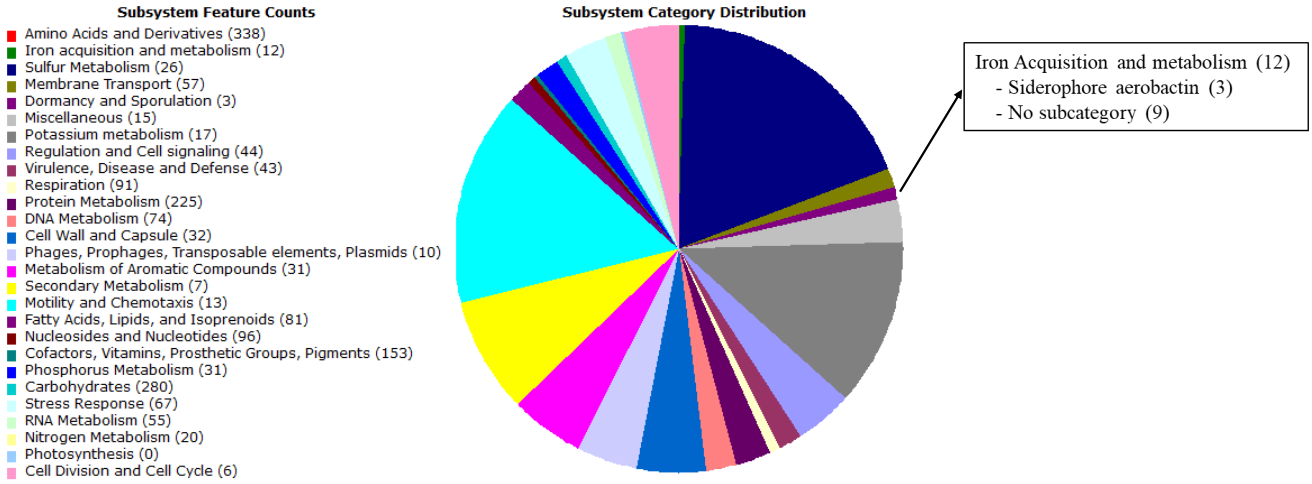

Figure 1. The results of RAST annotation of the M7 isolate

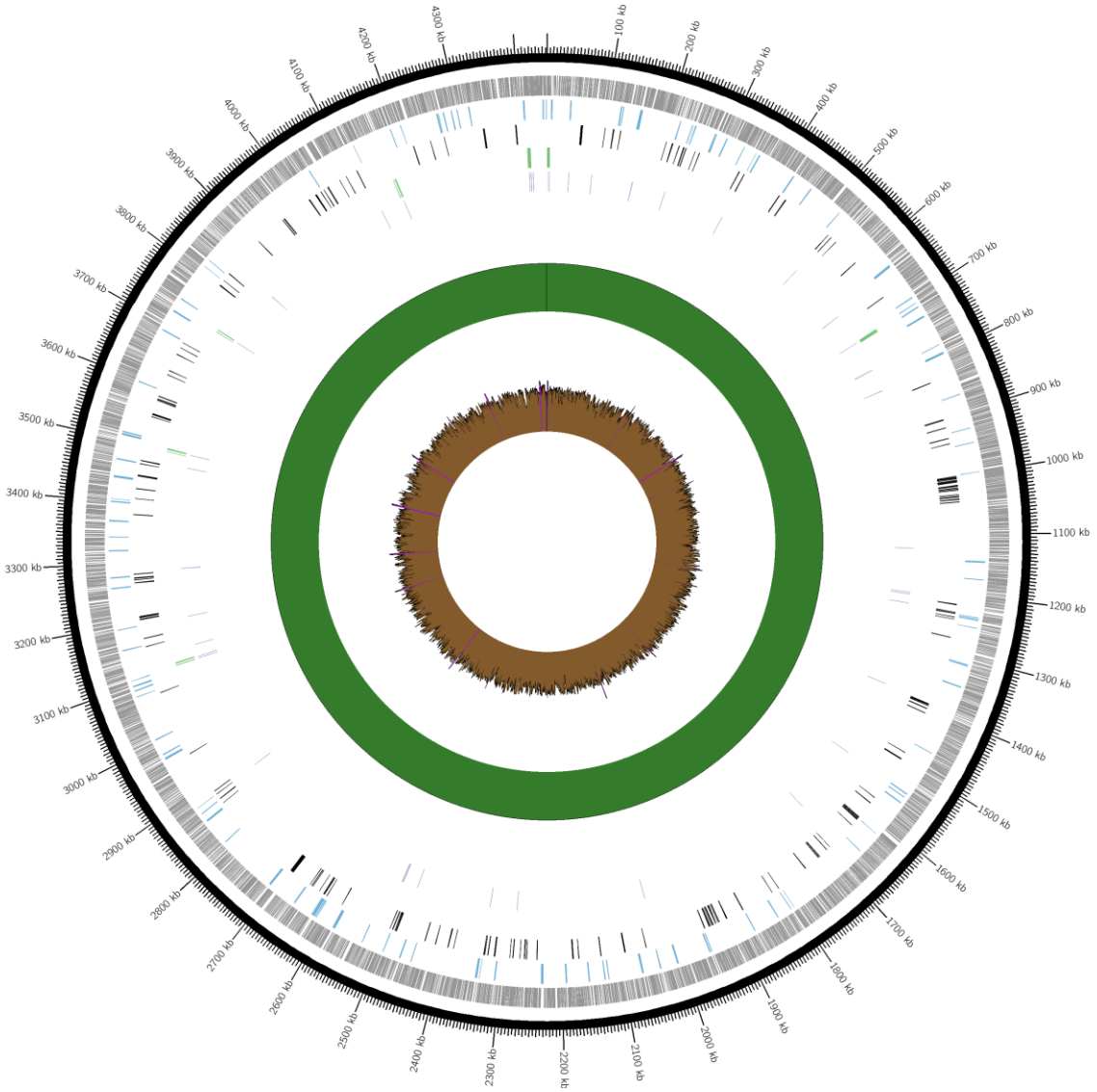

Figure 2. Circular maps of the M7 isolate genome created with Circos. Figure legend (from outermost to innermost): contig (black), genes (grey), pseudogenes (blue), CDS (black), rRNA (green), tRNA (purple), depth > 50 (green), depth < 50 (red), GC content > 50% (purple), GC content < 50% (brown)

Comparison of the M7 isolate genome

A comparative analysis of the ANI for the M7 isolate's genome was conducted with nine other *Providencia* strain genomes using the OrthoANI algorithm. The ANI analysis provides valuable insights into the phylogenetic relationships among the M7 isolate and other *Providencia* strains, contributing to a better understanding of the genetic diversity and taxonomic positioning of the M7 isolate within the *Providencia* genus. *Providencia* is a genus of Gram-negative bacteria within the order *Enterobacterales* and the family Morganellaceae. Pairwise ANI values between the M7 isolate and each *Providencia* strain, representing the genomes of recognized species, ranged from 76,67% to 98,82% (Figure 3). *Providencia manganoxydans*, *Providencia stuartii*, and *Providencia vermicola* are the three species which are most closely related to the M7 isolate, while *Providencia rettgeri* is the species with a more distant relationship to the M7 isolate. The M7 isolate exhibited an ANI of only 80,92% with the published strain *P. vermicola* P8538. In contrast, the M7 isolate showed high similarity with strain *P. manganoxydans* LLDRA6 and 36-1-2, with ANI values of 98.82% and 98.77%, respectively. These results indicate that the M7 isolate is closely related to *P. manganoxydans* species rather than *P. vermicola* species and *P. stuartii*, suggesting significant genetic and functional similarities with these closely related strains.

Siderophore genomic region in M7 isolate

According to the antiSMASH analysis for the bacterial genome, the genome of the M7 isolate and its three closest species contain distinct regions dedicated to secondary metabolite biosynthesis pathways (Table 2). These regions represent specific loci involved in producing various secondary metabolites, including siderophores. Siderophore biosynthesis is typically facilitated by NRPS and/or NIS pathways. In contrast to *P. vermicola*, which only contains biosynthesis gene clusters (BGCs) facilitated by the NRPS type, the M7 isolate, *P. manganoxydans*, and *P. stuartii* each have BGCs facilitated by both NRPS and NIS types.

Table 2. Comparison of secondary metabolite biosynthesis gene clusters in the M7 isolate with those of its closest strains

Clusters	Type	Location
M7 Isolate		
1	Thiopeptide	972,710-999,022 (total: 26,313 nt)
2	NRP-metallophore, NRPS	1,269,577-1,318,582 (total: 49,006 nt)
3	Betalactone	2,767,511-2,793,128 (total: 25,681 nt)
4	NIS	4,197,413-4,214,263 (total: 16,851 nt)
P. manganoxydans LLDRA6		
1	NIS	18,645-48,519 (total: 29,875 nt)
2	Betalactone	1,432,112-1,457,729 (total: 25,618 nt)
3	NRP-metallophore, NRPS	2,821,571-2,870,579 (total: 49,009 nt)
4	Thiopeptide	3,133,627-3,159,939 (total: 26,313 nt)
P. vermicola P8538		
1	NRP-metallophore, NRPS	916,932-992,935 (total: 76,004 nt)
2	Betalactone	1,371,296-1,396,783 (total: 25,488 nt)
3	Thiopeptide	3,144,069-3,170,421 (total: 26,353 nt)
P. stuartii FDAARGOS_645		
1	Betalactone	1,365,785-1,391,281 (total: 25,497 nt)
2	NRPS,T1PKS	1,821,746-1,877,078 (total: 55,333 nt)
3	NIS	2,790,977-2,820,833 (total: 29,857 nt)
4	Thiopeptide	4,035,199-4,061,444 (total: 26,246 nt)

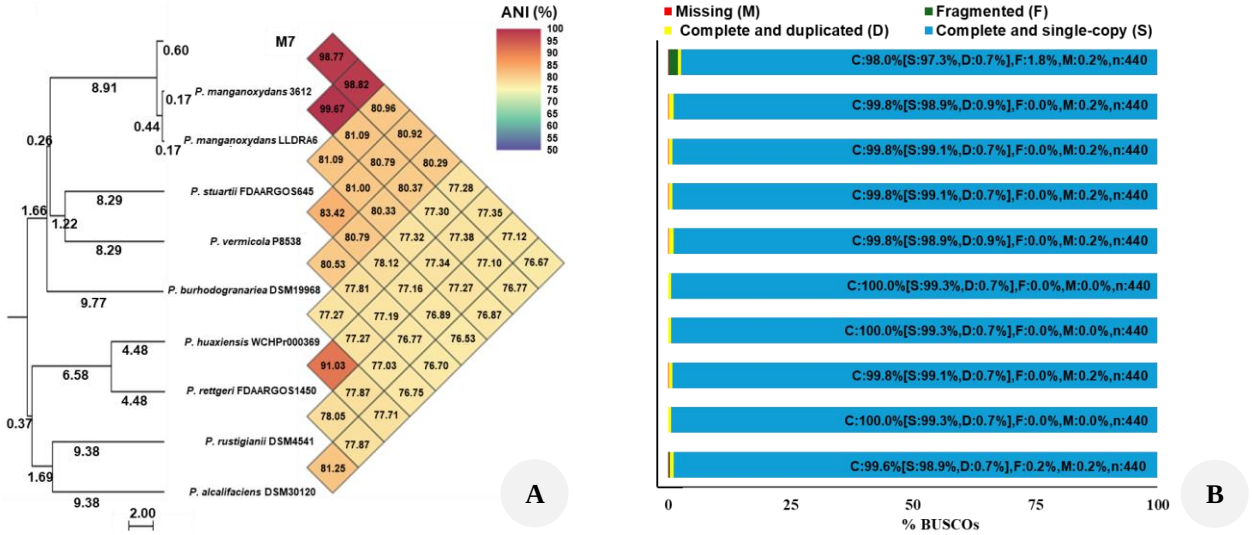

Figure 3. A. Phylogenetic tree of M7 isolate based on Average Nucleotides Identity (ANI) percentages comparison with other *Providencia* genomes using OAT software; B. The results of BUSCO assessments classified as complete single-copy (S, blue), complete duplicated (D, yellow), fragmented (F, green), and missing (M, red) for each genome

The cluster siderophore-dependent receptor content in each species was investigated to explore the potential clusters involved in siderophore biosynthesis. In Gram-negative bacteria, siderophore BGCs often include TonB-dependent siderophore receptors which are essential for the transport of siderophores across the bacterial membrane (Crits-Christoph et al. 2020). The cluster facilitated by the NIS pathway was identified as containing a TonB-dependent siderophore receptor in the M7 isolate, *P. manganoxydans*, and *P. stuartii*. However, in *P. vermicola*, this receptor is found in the cluster facilitated by the NRPS pathway. This difference in pathways results in varying mechanisms of iron acquisition.

Open reading frames (ORFs) related to siderophore biosynthesis within M7 isolate BGC #4 (Table 2), which contains a siderophore-dependent receptor in the M7 isolate, were examined in detail using the basic local alignment search tool for proteins (Blastp) (Figure 4). Furthermore, the six ORFs related to siderophore biosynthesis in cluster 4 exhibited significant similarities with genes from *P. stuartii* and *P. manganoxydans* (Table 3). Specifically, these homologous genes include those involved in the biosynthesis and regulation of siderophores, which are crucial for iron acquisition and microbial competitiveness. The substantial similarity between the ORFs in cluster 4 and those in these *Providencia* strains suggests that this cluster may be functionally analogous to known siderophore biosynthesis pathways in these *Providencia* species.

Comparison of siderophore biosynthesis gene clusters

The comparison of siderophore biosynthesis gene cluster structures from the M7 isolate with those from other genomes revealed a significant similarity of 94% with the cluster in *P. stuartii* (accession GCA_008693805.1) compared to its closest relative, *P. manganoxydans* (accession GCA_016618195.1) (Figure 5). Notably, the M7 isolate contains two distinct sequences of core biosynthesis genes crucial for siderophore production. In contrast, *P. stuartii* and *P. manganoxydans* each possess only a single core gene sequence responsible for siderophore biosynthesis. This structural disparity suggests that the M7 isolate may have a more complex or potentially more efficient biosynthetic pathway for siderophore production compared to these other strains.

Pathogenicity and virulence factors of the M7 isolate

To gain deeper insights into the properties of the M7 isolate, we conducted a series of analyses, including pathogenicity and virulence factor analysis, comparing the M7 isolate with *P. stuartii*, which has similar siderophore biosynthesis gene clusters (Table 4). The analysis using PathogenFinder indicates that M7 isolate has a lower potential to cause disease in humans compared to *P. stuartii*. This is evidenced by a lower probability of becoming a human pathogen for M7 compared to *P. stuartii*, as well as fewer pathogenic gene families associated with M7 compared to *P. stuartii*. The pathogenic gene families identified in the M7 isolate are not directly related to its pathogenic properties, whereas *P. stuartii* contains pathogenic genes that are 100% identical to those found in human pathogens such as *Proteus mirabilis* HI4320, *Salmonella enterica* subsp. *enterica* serovar Heidelberg str. SL476, *S. enterica* subsp. *enterica* serovar Typhi str. CT18, and *Escherichia fergusonii* str. ATCC 35469T. Nevertheless, analysis using VirulenceFinder did not identify specific virulence genes in either isolate.

Table 3. Predicted ORFs within cluster 4 and their top Blastp hits organism

ORF	Blast top hit organism	Query cover (%)	Identity (%)	Accession
1	*P. manganoxydans*	100	100	WP_349493617.1
2	*P. stuartii*	100	100	WP_070929816.1
3	*P. manganoxydans*	100	99.72	WP_369509851.1
4	*P. stuartii*	100	99.64	WP_140187320.1
5	*P. manganoxydans*	99	96.85	WP_369473596.1
6	*P. stuartii*	98	99.62	WP_349494436.1

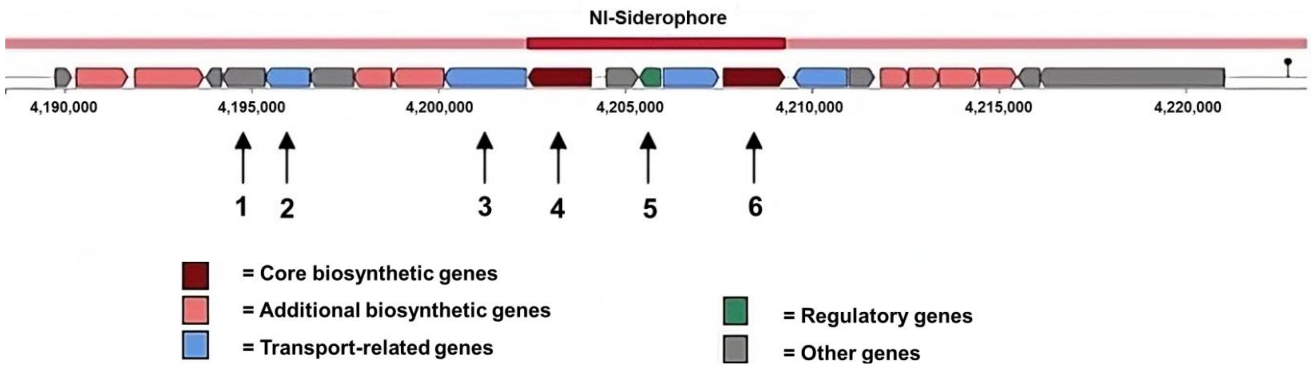

Figure 4. Prediction of ORFs within M7 isolate BGC #4 using antiSMASH. 1: ABC transporter; 2: MFS transporter; 3: TonB-dependent siderophore receptor; 4: IucA/IucC siderophore biosynthesis; 5: TetR/AcrR transcriptional regulator; 6: IucA/IucC siderophore biosynthesis

Table 4. Predicted pathogenicity and virulence factors of the M7 isolate compared with *P. stuartii*

Parameter	M7	*P. stuartii* FGAARGOS 645
Human pathogen	No	Yes
Probability of being a human pathogen	47.7 %	60.3 %
Proteome coverage	0.5 %	0.71 %
Matched pathogenic families	10	19
Matched not pathogenic families	10	9
VirulenceFinder predictions	Not hit found	Not hit found

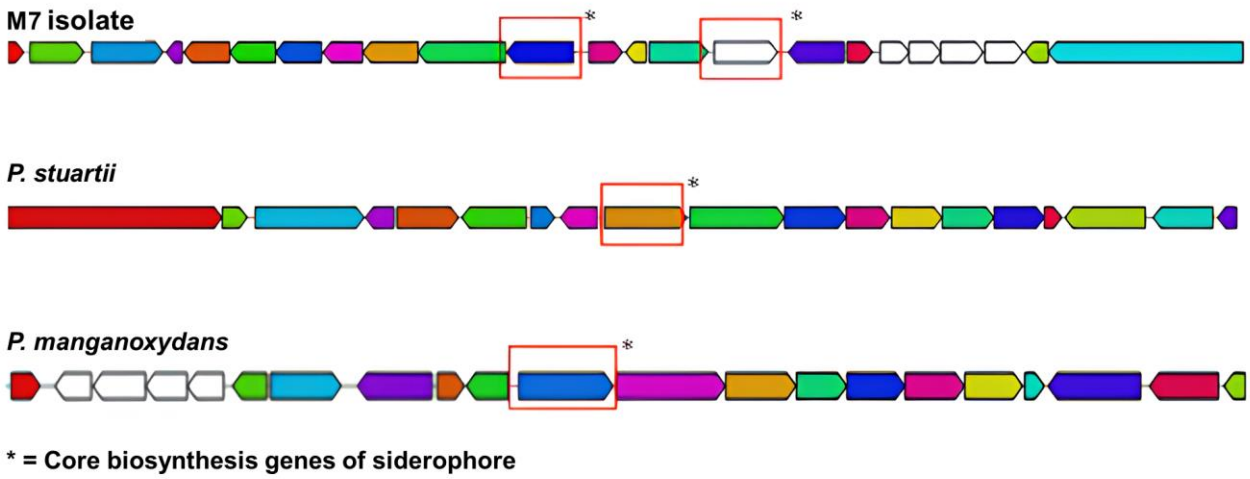

Figure 5. NIS BGC comparison of cluster 4 of the M7 isolate and similar cluster found in Blast top hit organism

Additionally, *P. manganoxydans*, a closely related bacterium to the M7 isolate, has not been implicated as a human pathogen in clinical settings. M7 isolate shares close genomic similarity with *P. manganoxydans*, suggesting a shared evolutionary history. The software analysis also utilizes a probability threshold of >50% to classify an organism as a potential human pathogen, and the probability of M7 isolate being a human pathogen was below this threshold, further suggesting its non-pathogenic nature (Cosentino et al. 2013).

Discussion

Siderophore systems are composed of low molecular weight molecules widespread across bacterial and fungal domains, with over 200 distinct biosynthetic types identified. These siderophores are crucial for iron uptake, a vital micronutrient, by capturing and solubilizing Fe(III) ions from the environment and facilitating their transport into the microbial cytoplasm (Li et al. 2016; Bruns et al. 2018). They achieve this by binding iron with high specificity and affinity, essential for organisms living in iron-limited environments.

The complete genome of the M7 isolate, a siderophore-producing bacterium isolated from Simadu pineapple roots in Subang District, West Java, Indonesia, was investigated. A GC depth analysis was performed to evaluate the potential for contamination during sequencing and assess the assembly's coverage, indicating that this genome has an average GC content of 40.18%, similar to the genus *Providencia* (Dong et al. 2024).

An evaluation of gene completeness using BUSCO shows a completeness level of 98% for the M7 isolate's genome, with fewer than ten genes being fragmented or missing. A high BUSCO gene completeness score typically reflects thorough genome annotation, suggesting that the sequencing data were of high quality and that the genome assembly was performed accurately, resulting in a well-constructed genome with minimal fragmentation (Manni et al. 2021). Additionally, a low level of fragmented genes implies that the genome annotation pipeline effectively captured the complexity of various gene models, accurately identifying and annotating most of the genes in the genome (Seppey et al. 2019).

In a previous study, the M7 isolate was predicted to be a *P. vermicola* species based on its 16S rRNA sequence (Fuadi et al. 2022). To further validate this identification, ANI values were used to compare the M7 isolate with other species. ANI analysis is particularly effective for distinguishing species within the same genus (Hussain et al. 2015; Du et al. 2022). Specifically, ANI values ranging from 95 to 96% signify species delineation among bacteria (Lee et al. 2016). These findings suggest that the M7 isolate is unrelated to *P. vermicola*. Instead, the ANI values indicate that the M7 isolate is more closely related to *P. manganoxydans* (Figure 3). *P. manganoxydans* was first isolated from heavy metal-contaminated soils in Hunan,

China. This strain has the potential to serve as a bio-adsorbent in environmental bioremediation to remove heavy metals like Pb, Cr, Cd, Cu, Mn, and Zn from contaminated soils (Li et al. 2020). However, no studies have yet explored siderophore biosynthesis in this strain. Additionally, siderophores are known to form complexes with heavy metals such as cadmium, lead, nickel, arsenic, aluminum, magnesium, zinc, copper, cobalt, and strontium (Sayyed and Chincholkar 2010; Sasirekha and Srividya 2016; Srivastava et al. 2022). This connection is further supported by an iron acquisition and metabolism system in the M7 isolate, which involves iron regulation mechanisms. Siderophore-dependent iron acquisition pathways, commonly found in various microbes (Miethke and Marahiel 2007; Segond et al. 2014; Li and Ma 2017; Li et al. 2019), suggest that the M7 isolate possesses a system for siderophore accumulation.

Studies have shown that the results obtained from antiSMASH analyses align well with laboratory findings used to identify siderophore BGCs (Lv et al. 2014; Cavas and Kirkiz 2022). The genome of the M7 isolate, which contains siderophore BGCs through the NIS pathway, was identified using antiSMASH. The NIS pathway involves a set of enzymatic processes responsible for synthesizing siderophores, critical for iron uptake in microbial organisms. In contrast, the identification of siderophore BGCs in the strain *P. vermicola* P8538 was facilitated by the NRPS pathway (Lupande-Mwenebitu et al. 2021). This pathway, distinct from the NIS pathway, also contributes to siderophores production but employs different enzymatic mechanisms and genetic clusters. The differences in the pathways involved in siderophore biosynthesis between *P. vermicola* and the M7 confirm previous studies, indicating that siderophore accumulation in the M7 isolate is higher than in *P. vermicola* (Fuadi et al. 2022). The higher siderophore accumulation in the M7 isolate suggests that the NIS pathway employed by the M7 isolate may offer enhanced efficiency or capacity for siderophore production. This increased efficacy could be attributed to the specific characteristics of the NIS pathway, such as its ability to produce siderophores in more significant quantities or with higher binding affinities for iron.

The compatibility of genes responsible for siderophore biosynthesis between the M7 isolate, *P. stuartii*, and *P. manganoxydans*, based on Blastp results, suggests that they may produce similar or related products. Additionally, two distinct nucleotide sequences serve as core biosynthesis genes, encoding the IucA/IucC family siderophore biosynthesis protein. This protein is responsible for aerobactin biosynthesis and catalyzes specific steps starting from N-epsilon-acetyl-N-epsilon-hydroxylysine and citrate (de Lorenzo and Neilands 1986). The presence of this gene duplication constitutes genetic redundancy, where two or more genes perform the same function. The presence of two core biosynthesis gene sequences in the M7 isolate may indicate a unique genetic adaptation or evolutionary advantage, potentially enhancing its ability to produce siderophores under diverse environmental conditions. Inactivating one copy of the gene has minimal to no effect on the biological traits essential for an organism's survival, as the other gene copies can still perform the necessary function (Nowak et al. 1997; Peng 2019). This allows both genes to express their proteins simultaneously. The increased presence of proteins responsible for siderophore biosynthesis concurrently enhances siderophore accumulation. This enhanced accumulation of siderophores could provide the M7 isolate with a competitive advantage in environments where iron acquisition is crucial. These findings suggest that while the siderophore biosynthesis structures of the M7 isolate and other strains are closely related, the presence of duplicated core genes in M7 may contribute to its higher siderophore accumulation, thus enhancing its ability to thrive in iron-limited environments.

The absence of detected virulence genes in the M7 isolate suggests that it does not carry genes known to contribute to pathogenicity, according to the criteria used by the tool (Joensen et al. 2014). The non-pathogenic status of M7 isolate is further supported by its low probability of being classified as a human pathogen, considering the absence of virulence factors and pathogenicity islands. In contrast, *P. stuartii* FGAARGOS 645 harbors a pathogenic gene family, including a tetracycline resistance protein similar to that found in *E. fergusonii* ATCC 35469T. This protein confers tetracycline resistance by actively effluxing the antibiotic as a metal-tetracycline/H+ antiporter, thereby reducing its accumulation in bacterial cells. Additionally, studies have shown that *P. stuartii* contains multiple virulence genes associated with pathogenicity, such as *imA*, *mrkA*, *fptA*, *iutA*, *ireA*, and *hlyA*, and is known to produce hemolysin (Guidone et al. 2023). Although virulence gene families in M7 isolate were not detected, and *P. manganoxydans*-the closest relative of M7 isolate-has not been reported to cause human disease, further research is needed. This includes in vivo experiments, such as animal testing, to accurately assess the potential risks posed by M7 isolate. These findings are preliminary, and experimental validation will be essential to confirm the pathogenic potential of the M7 isolate.

In conclusion, the genome of the new siderophore-producing bacterium is consistent with that of the *P. manganoxydans* species and closely related to the strain *P. manganoxydans* LLDRA6. However, regarding siderophore BGCs, this isolate shows the highest similarity to *P. stuartii*, with notable differences in genetic redundancy. These differences allow the bacterium to achieve high levels of siderophore biosynthesis through the NIS pathway, utilizing two core siderophore biosynthesis genes. These duplicated core genes likely enhance its siderophore production capacity, providing a competitive advantage in environments where iron acquisition is essential.

ACKNOWLEDGEMENTS

We would like to acknowledge Institut Pertanian Bogor, Indonesia, for providing financial support through the Fundamental Research Grant Program (No. 445/IT3.D10/ PT.01.03/P/B/2023). We also extend our gratitude to all those who contributed to the successful completion of this research.

REFERENCES

Abdallah M, Balshi A. 2018. First literature review of carbapenem-resistant *Providencia*. New Microb New Infect 25: 16-23. DOI: 10.1016/j.nmni.2018.05.009.

Ameer AKA, Al-Gburi NM. 2022. Study the pathogenicity of *Providencia vermicola* in mice. Intl J Health Sci 6: 6895-6906. DOI: 10.53730/ijhs.v6nS4.10020.

Baars O, Zhang X, Morel FMM, Seyedsayamdost MR. 2016. The Siderophore metabolome of *Azotobacter vinelandii*. Appl Environ Microbiol 82 (1): 27-39. DOI: 10.1128/AEM.03160-15.

Bajbouj K, Shafarin J, Hamad M. 2018. High-dose deferoxamine treatment disrupts intracellular iron homeostasis, reduces growth, and induces apoptosis in metastatic and nonmetastatic breast cancer cell lines. Technol Cancer Res Treat 17: 1533033818764470. DOI: 10.1177/1533033818764470.

Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T. 2023. antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures, and visualization. Nucleic Acids Res 51 (W1): W46-W50. DOI: 10.1093/nar/gkad344.

Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucl Acids Res 49 (W1): W29-W35. DOI: 10.1093/nar/gkab335.

Bosshard PP, Abels S, Zbinden R, Böttger EC, Altwegg M. 2003. Ribosomal DNA sequencing for identification of aerobic Gram-positive rods in the clinical laboratory (an 18-Month Evaluation). J Clin Microbiol 41 (9): 4134-4140. DOI: 10.1128/JCM.41.9.4134-4140.2003.

Bradley JM, Svistunenko DA, Wilson MT, Hemmings AM, Moore GR, Le Brun NE. 2020. Bacterial iron detoxification at the molecular level. J Biol Chem 295: 17602-17623. DOI: 10.1074/jbc.REV120.007746.

Bruns H, Crüsemann M, Letzel A-C, Alanjary M, McInerney JO, Jensen PR, Schulz S, Moore BS, Ziemert N. 2018. Function-related replacement of bacterial siderophore pathways. Intl Soc Microb Ecol 12 (2): 320-329. DOI: 10.1038/ismej.2017.137.

Cavas L, Kirkiz I. 2022. Characterization of siderophores from *Escherichia coli* strains through genome mining tools: An antiSMASH study. Appl Microbiol Biotechnol Expr 12 (1): 74-87. DOI: 10.1186/s13568-022-01421-x.

Chandrangsu P, Rensing C, Helmann JD. 2017. Metal homeostasis and resistance in bacteria. Nat Rev Microbiol 15 (6): 338-350. DOI: 10.1038/nrmicro.2017.15.

Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. 2013. PathogenFinder-distinguishing friend from foe using bacterial whole genome sequence data. PLoS One 8 (10): e77302. DOI: 10.1371/journal.pone.0077302.

Crits-Christoph A, Bhattacharya N, Olm M, Song Y, Banfield J. 2020. Transporter genes in biosynthetic gene clusters predict metabolite characteristics and siderophore activity. Genome Res 31 (2): 239-250. DOI: 10.1101%2Fgr.268169.120.

Cui Y, Cui Z, Xu J, Hao D, Shi J, Wang D, Xiao H, Duan X, Chen R, Li W. 2020. NG-Circos: Next-generation Circos for data visualization and interpretation. NAR Genom Bioinform 2 (3): lqaa069. DOI: 10.1093/nargab/lqaa069.

De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: Visualizing and processing long-read sequencing data. Bioinformatics 34: 2666-2669. DOI: 10.1093/bioinformatics/bty149.

de Lorenzo V, Neilands JB. 1986. Characterization of *iucA* and *iucC* genes of the aerobactin system of plasmid ColV-K30 in *Escherichia coli*. J Bacteriol 167: 350-355. DOI: 10.1128/jb.167.1.350-355.1986.

Dong X, Jia H, Yu Y, Xiang Y, Zhang Y. 2024. Genomic revisitation and reclassification of the genus *Providencia*. mSphere 9 (3): e00731-23. DOI: 10.1128/msphere.00731-23.

Drechsel H, Jung G, Winkelmann G. 1993. Alpha-keto acids are novel siderophores in the genera *Proteus*, *Providencia*, and *Morganella* and are produced by amino acid deaminases. J Bacteriol 175 (9): 2727-2733. DOI: 10.1128/jb.175.9.2727-2733.1993.

Du Y, Li X, Liu Y, Mu S, Shen D, Fan S, Lou Z, Zhang S, Xia H, Yuan Y, Wang S. 2022. The species identification and genomic analysis of *Haemobacillus shengwangii*: A novel pathogenic bacterium isolated from a critically ill patient with bloodstream infection. Front Microbiol 13: 919169. DOI: 10.3389/fmicb.2022.919169.

Fuadi H, Suryadarma P, Syamsu K, Surono S, Setiyani NA, Ridhoha SM, Zahra AS, Stepani N, Ramadhan MR. 2022. Isolation and selection of siderophore-producing bacteria from roots of Simadu pineapple (*Ananas comosus*) in Subang District, West Java. Menara Perkebunan 90 (2): 160-167. DOI: 10.22302/iribb.jur.mpv90i2.502.

Gu S, Shao Y, Rehm K, Bigler L, Zhang D, He R, Xu R, Shao J, Jousset A, Friman VP, Bian X, Wei Z, Kümmerli L, Li Z. 2024. Feature sequence-based genome mining uncovers the hidden diversity of bacterial siderophore pathways. eLife 13: RP96719. DOI: 10.7554/eLife.96719.3.

Guidone GHM, Cardozo JG, Silva LC, Sanches MS, Galhardi LCF, Kobayashi RKT, Vespero EC, Rocha SPD. 2023. Epidemiology and characterization of *Providencia stuartii* isolated from hospitalized patients in southern Brazil: A possible emerging pathogen. MicroSoc 55 (10): 000652-v4. DOI: 10.1099/acmi.0.000652.v2.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29 (8): 1072-1075. DOI: 10.1093/bioinformatics/btt086.

Ho Y-H, Ho S-Y, Hsu C-C, Shie J-J, Wang T-SA. 2017. Utilizing an iron(III)-chelation masking strategy to prepare mono- and bis-functionalized aerobactin analogues for targeting pathogenic bacteria. Chem Commun 53 (66): 9265-9268. DOI: 10.1039/C7CC05197B.

Hussain K, Hameed S, Shahid M, Ali A, Iqbal J, Hahn D. 2015. First report of *Providencia vermicola* strains characterized for enhanced rapeseed growth attributing parameters. Intl J Agric Biol 17 (6): 1110-1116. DOI: 10.17957/IJAB/15.0033.

Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol 52 (5): 1501-1510. DOI: 10.1128/JCM.03617-13.

Klein JA, Predeus AV, Greissl AR, Clark-Herrera MM, Cruz E, Cundiff JA, Haeberle AL, Howell M, Lele A, Robinson DJ, Westerman TL, Wrande M, Wright SJ, Green NM, Vallance BA, McClelland M, Mejia A, Goodman AG, Elfenbein JR, Knodler LA. 2024. Pathogenic diversification of the gut commensal *Providencia alcalifaciens* via acquisition of a second type III secretion system. Infect Immun 92 (10): e00314-24. DOI: 10.1128/iai.00314-24.

Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17 (11): 1103-1110. DOI: 10.1038/s41592-020-00971-x.

Lee I, Ouk Kim Y, Park S-C, Chun J. 2016. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. Intl J Syst Evol Microbiol 66: 1100-1103. DOI: 10.1099/ijsem.0.000760.

Li C, Zhu L, Pan D, Li S, Xiao H, Zhang Z, Shen X, Wang Y, Long M. 2019. Siderophore-mediated iron acquisition enhances resistance to oxidative and aromatic compound stress in *Cupriavidus necator* JMP134. Appl Environ Microbiol 85 (1): e01938-18. DOI: 10.1128/AEM.0193818.

Li D, Li R, Ding Z, Ruan X, Luo J, Chen J, Zheng J, Tang J. 2020. Discovery of a novel native bacterium of *Providencia* sp. with high biosorption and oxidation ability of manganese for bioleaching of heavy metal contaminated soils. Chemosphere 241: 125039. DOI: 10.1016/j.chemosphere.2019.125039.

Li K, Chen W-H, Bruner SD. 2016. Microbial siderophore-based iron assimilation and therapeutic applications. Biometals 29: 377-388. DOI: 10.1007/s10534-016-9935-3.

Li Y, Ma Q. 2017. Iron acquisition strategies of *Vibrio anguillarum*. Front Cell Infect Microbiol 7: 342-354. DOI: 10.3389/fcimb.2017.00342.

Lupande-Mwenebitu D, Khedher MB, Khabthani S, Rym L, Phoba MF, Nabti LZ, Lunguya-Metila O, Pantel A, Lavigne JP, Rolain JM, Diene SM. 2021. First genome description of *Providencia vermicola* isolate bearing NDM-1 from blood culture. Microorganisms 9 (8): 1751. DOI: 10.3390/microorganisms9081751.

Lv H, Hung CS, Henderson JP. 2014. Metabolomic analysis of siderophore cheater mutants reveals metabolic costs of expression in uropathogenic *Escherichia coli*. J Proteome Res 13 (3): 1397-1404. DOI: 10.1021/pr4009749.

Majewska M, Słomka A, Hanaka A. 2024. Siderophore-producing bacteria from Spitsbergen soils-novel agents assisted in bioremediation of the metal-polluted soils. Environ Sci Pollut Res 31: 32371-32381. DOI: 10.1007/s11356-024-33356-0.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM, Joana K. 2021. BUSCO Update: Novel and streamlined workflows along with

broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol 38 (10): 4647-4654. DOI: 10.1093/molbev/msab199.

Medema MH, Blin K, Cimermancic P, De Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. 2011. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res 39: W339-W346. DOI: 10.1093/nar/gkr466.

Miethke M, Marahiel MA. 2007. Siderophore-based iron acquisition and pathogen control. Microbiol Mol Biol Rev 71 (3): 413-451. DOI: 10.1128/MMBR.00012-07.

Nowak MA, Boerlijst MC, Cooke J, Smith JM. 1997. Evolution of genetic redundancy. Nature 388 (6638): 167-171. DOI: 10.1038/40618.

Paredes C, Chi SI, Flint A, Weedmark K, McDonald C, Bearne J, Ramirez-Arcos S, Pagotto F. 2021. Complete genome sequence of *Staphylococcus aureus* PS/BAC/169/17/W, isolated from a contaminated platelet concentrate in england. Microbiol Resour Announc 10 (45): e00841-21. DOI: 10.1128/MRA.00841-21.

Peng J. 2019. Gene redundancy and gene compensation: An updated view. J Genet Genom 46 (7): 329-333. DOI: 10.1016/j.jgg.2019.07.001.

Saha M, Sarkar S, Sarkar B, Sharma BK, Bhattacharjee S, Tribedi P. 2016. Microbial siderophores and their potential applications: A review. Environ Sci Pollut Res 23: 3984-3999. DOI: 10.1007/s11356-015-4294-0.

Sasirekha B, Srividya S. 2016. Siderophore production by *Pseudomonas aeruginosa* FP6, a biocontrol strain for *Rhizoctonia solani* and *Colletotrichum gloeosporioides* causing diseases in chili. Agric Nat Resour 50 (4): 250-256. DOI: 10.1016/j.anres.2016.02.003.

Sayyed RZ, Chincholkar SB. 2010. Growth and siderophores production in *Alcaligenes faecalis* is regulated by metal ions. Indian J Microbiol 50: 179-182. DOI: 10.1007/s12088-010-0021-1.

Segond D, Abi Khalil E, Buisson C, Daou N, Kallassy M, Lereclus D, Arosio P, Bou-Abdallah F, Nielsen Le Roux C. 2014. Iron acquisition in *Bacillus cereus*: The roles of *IlsA* and bacillibactin in exogenous ferritin iron mobilization. PLoS Pathog 10 (2): e1003935. DOI: 10.1371/journal.ppat.1003935.

Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: Assessing genome assembly and annotation completeness. Gene Prediction 1962: 227-25. DOI: 10.1007/978-1-4939-9173-0_14.

Seyoum Y, Baye K, Humblot C. 2021. Iron homeostasis in host and gut bacteria - a complex interrelationship. Gut Microbes 13 (1): 1874855. DOI: 10.1080/19490976.2021.1874855.

Simao FA, Waterhouse RM, Ionnidis P, Kriventseva EV, Zdobnov EM, 2015. BUSCO: Assessing genome assembly and annotation completeness with single copy orthologs. Bioinformatics 31 (19): 3210-3212. DOI: 10.1093/bioinformatics/btv351.

Srivastava P, Sahgal M, Sharma K, Enshasy HAE, Gafur A, Alfarraj S, Ansari MJ, Sayyed RZ. 2022. Optimization and identification of siderophores produced by *Pseudomonas monteilii* strain MN759447 and its antagonism toward fungi associated with mortality in *Dalbergia sissoo* plantation forests. Front Plant Sci 13: 984522. DOI: 10.3389/fpls.2022.984522.

Swayambhu G, Bruno M, Gulick AM, Pfeifer BA. 2021. Siderophore natural products as pharmaceutical agents. Curr Opin Biotechnol 69: 242-251. DOI: 10.1016/j.copbio.2021.01.021.

Tamang P, Upadhaya A, Paudel P, Meepagala K, Cantrell CL. 2024. Mining biosynthetic gene clusters of *Pseudomonas vancouverensis* utilizing whole genome sequencing. Microorganisms 12 (3): 548-565. DOI: 10.3390/microorganisms12030548.

Telfer TJ, Liddell JR, Duncan C, White AR, Codd R. 2017. Adamantyl- and other polycyclic cage-based conjugates of desferrioxamine B (DFOB) for treating iron-mediated toxicity in cell models of parkinson's disease. Bioorg Med Chem Lett 27 (8): 1698-1704. DOI: 10.1016/j.bmcl.2017.03.001.

Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27 (5): 737-746. DOI: 10.1101/gr.214270.116.

Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43 (W1): W237-W243. DOI: 10.1093/nar/gkv437.

Wilson BR, Bogdan AR, Miyazawa M, Hashimoto K, Tsuji Y. 2016. Siderophores in iron metabolism: From mechanism to therapy potential. Trends Mol Med 22: 1077-1090. DOI: 10.1016/j.molmed.2016.10.005.

Yu S, Teng C, Bai X, Liang J, Song T, Dong L, Jin Y, Qu J. 2017. Optimization of siderophore production by *Bacillus* sp. PZ-1 and its potential enhancement of phytoextration of pb from soil. J Microbiol Biotechnol 27 (8): 1500-1512. DOI: 10.4014/jmb.1705.05021.

Zavascki AP, Goldani LZ, Li J, Nation RL. 2007. Polymyxin B for the treatment of multidrug-resistant pathogens: A critical review. J Antimicrob Chemother 60 (6): 1206-1215. DOI: 10.1093/jac/dkm357.